

12-3-96

EXHIBIT B

Bali Floating Point Representation

This document defines the 16-bit floating point format stored in pixel and texel memory known as "s10e5". Internal variations used inside the R chip pipeline known as "s11e5" and "s15e5" have 11 or 15 mantissa bits instead of the 10 shown here.

Readers may notice some similarity to the IEEE floating point standard :-)



Value	Conditions
$(-1)^s \times 2^{(e-15)} \times 1.m$	$0 < e < 31$
zero	$e == 0$, all values of m and s
positive infinity	$e == 31$, $s == 0$, all values of m
negative infinity	$e == 31$, $s == 1$, all values of m

excess 15 (bias = 15)
 $\beta = 2$

As an example, the value (-1.25) is represented by $s=1$, $m=0100000000$ (binary), $e=01111$ (binary).

Note that denormalized numbers are not supported. Underflows are forced to zero; overflows are clamped to positive or negative infinity.

Have to convert from IEEE 32 bit float to this in convert/merge block.